1. Definimos las variables aleatorias:

$$\begin{aligned} D|E &= \textit{distancia a la que Hulk lanza un coche enfadado} \sim N(700, 150) \\ D|\bar{E} &= \textit{distancia a la que Hulk lanza el coche no enfadado} \sim N(500, 100) \\ D &= \textit{distancia a la que lanza el coche} \end{aligned}$$

   Conocemos la probabilidad de que esté enfadado, $P(E) = 0.8$ y $P(\bar{E}) = 0.2$. Se ha observado que ha lanzado el coche a más de 600 metros y se pide la probabilidad de que no esté enfadado $P(\bar{E}|D > 600)$. Antes de usar el teorema de Bayes para calcularla, hay que calcular las probabilidades condicionadas y usar el Teorema de la Probabilidad Total:

$$P(D > 600) = P(D > 600|E)P(E) + P(D > 600|\bar{E})P(\bar{E})$$

   Y para las probabilidades condicionadas usamos la distribución Normal:

$$\begin{aligned} P(D > 600|E) &= P\left(Z > \frac{600 - 700}{150}\right) \\ &= P(Z > -0.66) = P(Z < 0.66) = 0.7454 \end{aligned}$$

$$\begin{aligned} P(D > 600|\bar{E}) &= P\left(Z > \frac{600 - 500}{100}\right) \\ &= P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1587 \end{aligned}$$

   Y, por tanto,

$$P(D > 600) = (0.7454) \cdot (0.8) + (0.1587) \cdot (0.2) = 0.628$$

   Y se usa el Teorema de Bayes para calcular la probabilidad pedida:

$$P(\bar{E}|D > 600) = \frac{P(D > 600|\bar{E})P(\bar{E})}{P(D > 600)} = \frac{0.0317}{0.628} = 0.05$$

2. El tiempo hasta fallo en las cámaras se distribuye como Exponencial de media 25000 h., por tanto, de parámetro $\lambda = \frac{1}{25000}$. Nos piden:

   a) $P(T > 20000) = 1 - P(T \leq 20000) = e^{-\frac{20000}{25000}} = 0.4493$

   b) $P(T < 30000) = 1 - e^{-\frac{30000}{25000}} = 0.6988$

   c) $P(25000 < T < 30000) = P(T < 30000) - P(T < 25000) = 0.6988 - (1 - 0.4493) = 0.1481$

d) Como la varianza de la exponencial es $\frac{1}{\lambda^2}$, su desviación típica es $\frac{1}{\lambda} = 25000$. Por tanto, la probabilidad pedida es:

$$P(T > 2 \cdot (25000)) = P(T > 50000) = e^{-\frac{50000}{25000}} = e^{-2} = 0.1353$$

Nota: Las probabilidades se calculan con la ***función de distribución*** *de la exponencial que* ***no*** *es la que viene en el formulario. En el formulario viene la función de densidad, integrándola desde 0 a t se obtiene la función de distribución $F(t) = P(T \leq t) = 1 - e^{-\lambda t}$. Se pide $P(T > t) = 1 - P(T \leq t) = e^{-\lambda t}$. También puede usarse la función de distribución de la Erlang con λ y $p = 1$.*

3. Se pide un intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$, suponiendo varianzas desconocidas pero iguales. La variable pivote será:

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

con lo que el IC al $(1-\alpha)100\,\%$ es:

$$\left(\bar{X} - \bar{Y} - t_{n_1+n_2-2,\frac{\alpha}{2}} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \bar{X} - \bar{Y} + t_{n_1+n_2-2,\frac{\alpha}{2}} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$$

Hay que hacer algunos cálculos:

$$s_1^2 = \frac{157398.2 - 7(149.95)^2}{6} = 0.53, \qquad s_2^2 = \frac{139104 - 6(152.26)^2}{5} = 1.07$$

y a partir de aquí:

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{6 \times 0.53 + 5 \times 1.07}{11} = 0.775, \quad s_p = 0.88$$

Se necesita también el percentil $t_{11,0.05} = 1.796$ ya que el intervalo es al $90\,\%$, $1-\alpha = 0.9$. El intervalo es:

$$-2.31 \pm (1.796)(0.88)\sqrt{\frac{1}{7} + \frac{1}{6}}$$

$$\mu_1 - \mu_2 \in (-3.189, -1.431) \text{ con una confianza del } 90\,\%$$

A la vista del intervalo, como es negativo, se puede afirmar con un $90\,\%$ de confianza que las chicas cobran menos que los chicos ($\mu_1 < \mu_2$), como María suponía.

4. Contrastamos la hipótesis de que la varianza es igual a 4 frente a la alternativa de que es distinta (contraste bilateral). No conocemos la media, por tanto la estimamos a partir de los datos y:

$$\left\{ \begin{array}{lcl} H_0 & : & \sigma_1^2 = 4 \\ H_1 & : & \sigma_1^2 \neq 4 \end{array} \right. \qquad \text{Medida de discrepancia} \quad \frac{(n-1)s^2}{\sigma^2} \underbrace{\sim}_{\text{bajo } H_0} \chi^2_{n-1}$$

A partir de los datos que se dan, se calcula lo que necesitamos:

$$\sum_{i=1}^{10} x_i = 76.2, \;\; \bar{x} = \frac{76.2}{10} = 7.62$$

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \frac{590.32 - 10(7.62)^2}{9} = 1.075$$

Se calcula $\hat{d}$,

$$\hat{d} = \frac{(n-1)s^2}{\sigma^2} = \frac{9(1.075)}{4} = 2.41$$

La región de no rechazo del test es, para $\alpha = 0.02$:

$$(\chi^2_{9,0.99}, \chi^2_{9,0.01}) = (2.088, 21.666)$$

Como $\hat{d}$ pertenece a esa región de no rechazo no tenemos evidencias para rechazar H_0 para un nivel de significación de 0.02 y la varianza sí puede suponerse igual a 4 y la desviación típica igual a 2.